

Case Study: BlueCrest Capital Management Ltd.

Real-Time Market Data Management

By Bob Giffords
Independent Banking and Technology Analyst

Motivation for Real-Time Market Data Management

BlueCrest Capital Management is a leading European hedge fund, based in London, with over \$15 billion in assets-under-management and an award-winning reputation for combining 'technology, diligence and innovation to produce the most efficient risk-taking possible'.

Since its launch in 2000, BlueCrest has aimed to stay ahead of the curve by anticipating future challenges and investing early in skills and technology. The firm's traders are highly technical with many trained mathematicians. Complex statistical models form the foundation of some of its trading strategies, and these rely on a rich variety of data-feeds to enable them to trade "anything that ticks", as their Head of Trading Strategy Systems, Justo Ruiz-Ferrer describes it, adding, "We trade everything from traditional fixed income, equities, FX and related derivatives, to exotic commodity, energy and alternative asset classes." In such highly competitive markets, BlueCrest soon realized that good data management was as important as their quantitative models. "You may have a Ferrari, but if you fill up your tank at the supermarket, you won't get the performance," observes Ruiz-Ferrer.

Scalability and real-time performance were key factors with BlueCrest's data storage demands exploding at 200 to 400% per year, and exchange based derivative markets hitting new highs of billions of price ticks per day, not to mention the OTC platforms, as ever more robotraders enter the jittery markets.

A New Market Data Management System

"BlueCrest has ensured that its vital market data gets the highest octane data mix available to stay ahead of the pack and negotiate the market chicanes at full throttle."

**Justo Ruiz-Ferrer,
Head of Trading
Strategy Systems,
BlueCrest Capital
Management**

In 2007, just as the credit crisis was breaking, BlueCrest set up a team under Justo Ruiz-Ferrer, to develop a state-of-the-art market data management system. BlueCrest trades 24 hours a day, six days a week, across multiple markets using a wide range of data feeds. As markets move day to day and week to week, BlueCrest needed to rapidly reconfigure data feed connections and plug the data into real-time models while optimizing management of the necessary data feed licenses. BlueCrest devised a solution that combines the rapid time-to-market event processing capabilities of StreamBase with the instant storage and retrieval functionality of Vertica. It provides a total market data management solution that is able to meet the needs of low-latency trading and the demanding innovation of their quantitative analysts to achieve greater profitability. As the range of data sources constantly grows with increased cross asset class and exotic derivatives, time to market, low latency and auditable compliance with complex data licensing rules were essential.

Ad hoc engineering of new data feed adapters had proven to be both expensive and risky due to complex software interfaces, few standards, and low reusability of components. The new system therefore aimed to decouple users from the data sources to create a common interface for accessing, conflating, enriching and distributing real-time data with minimal latency and assured compliance. Yet it also had to encourage flexible data sharing across the firm.

BlueCrest's new market data system provides both current and historical snapshots and on-demand event processing for both simple requests like last price traded to complex scenarios like volatility surfaces or correlation signals with time windows. It also has to guide the user through the rich jungle of data sources and compliance rules, and help resolve queries that may have alternative solutions across the data providers. Each source and derived data-object needs its own access controls, as these cannot always be inferred.

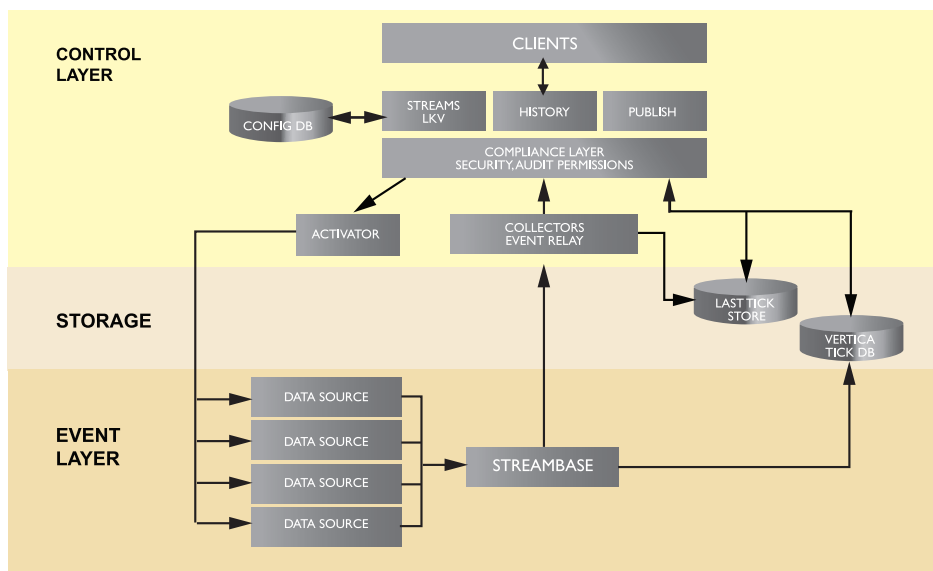
Market Data Mechanics

The system was built in three layers: (1) An event layer that includes the data sources and a core event-processing engine from StreamBase to extract, collate and enrich the source data; (2) The storage layer consists of a real-time, historical tick database from Vertica, backed up on disc, and a last tick store for last known value (LKV) caching; while (3) the control layer includes compliance and management components written in-house in C#. The control layer checks user access permissions, retrieves data from the storage layer or activates data streams and then dynamically sets up 'Collectors' to relay the resulting 'event-curves', such as volatility, yield or momentum curves to all subscribed users or their software applications. The whole system is driven by real-time requests from the users or their software models.

BlueCrest's 3 layered-system for market data management consists of:

- (1) Event layer
- (2) Storage layer
- (3) Control layer

Data streams fuel the system. Ticker plant and internal messaging buses feed data to the StreamBase complex event processing engine to clean, normalize, filter and enrich the event-curves. As part of this enrichment,



StreamBase can mix real-time events with the historical data in Vertica as needed to evaluate the complex rules, yet still keep pace with the input data rates. StreamBase then forwards the new event curve objects to both Vertica for storage and the Collectors in the control layer, which encrypt them to ensure compliance, cache them in the last tick store, and distribute them to all waiting user processes. Whenever any point changes, StreamBase generates a new event, and therefore provides a stream of event curves rather than just data points. "It's immensely powerful given the high speed inputs and the complex handling requirements," says Ruiz-Ferrer.

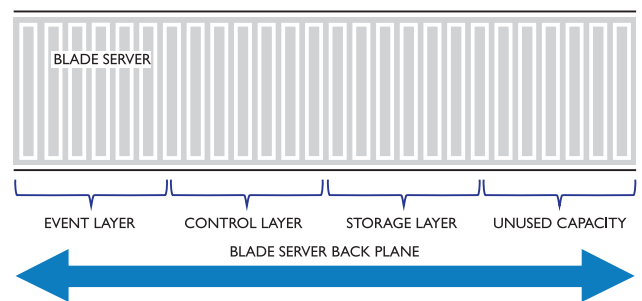
The system works round the clock six days a week supporting user services across different geographical locations. Further servers with collectors, local caches and control software are also located in each office to minimize latencies for last value type queries, which represent the bulk of the workload.

All the core data feeds are now plumbed into the system including the major data aggregators and many direct exchange feeds. However, much of the data actually comes from internal BlueCrest services over the messaging bus. "We can subscribe to everything," says Ruiz-Ferrer, "so we load tested to ten times the whole feed, just to be safe."

Security is tight, hence the encryption, which is the most expensive component in terms of CPU resource, according to Ruiz-Ferrer. Moreover, the system keeps track of who is accessing which service at any time to ensure compliance with the licensing rules. One advantage is that BlueCrest can optimize license fees to maximize sharing when data is not in use. “What a user sees depends on the market data licence. Without the licence, users only see limited price movements, whereas with the license they see everything,” explains Ruiz-Ferrer.

In London all the processes run on a single blade server to eliminate network hops. Each layer of the architecture can be separately scaled across its own set of server blades to meet changing performance and resilience needs. Thus the whole sequence from real-time source to waiting user servers in London takes under 20 milliseconds on a benchmark test of 2 million updates per second with very low server loadings.

For the middle office and research models, the priority was to leverage the full knowledge base of the firm with great flexibility and rich functionality rather than minimizing microseconds. Indeed the core StreamBase component can typically be completed in well under a millisecond, but complex enrichment and security do add some latency.



Rapid Time to Market

The project build started in January and by May it was in production, even allowing for a one-month delay due to other priorities. The team now is working to port other data feeds and analytics to StreamBase, enriching the visualization tools, as well as making the whole system fail-safe with a high speed interconnect to the back-up site.

Prior to the market data management solution it could take 2 months or more to add and integrate a new data feed. Now a new service is connected in just two weeks because everything is standardized. “This makes the whole business much more responsive and confident,” says Ruiz-Ferrer.

Having evaluated many other solutions, the BlueCrest team chose StreamBase because of its real-time performance and flexibility, with many adapters available out-of-the-box for connecting to data feed handlers, messaging and data base systems including Vertica, and trading platforms. They found the visual workflow-like design capabilities of StreamBase particularly helpful for rapid development and maintenance.

“StreamBase and Vertica have allowed us to develop our market data environment rapidly with a sound run time behaviour, excellent extensibility and minimum effort”, concludes Ruiz-Ferrer. “By investing in a high-speed data storage and an ultra low-latency complex event processing solution, BlueCrest has ensured that its vital market models get the highest octane data mix available to stay ahead of the pack and negotiate the market chicanes at full throttle.” BlueCrest has demonstrated once again how innovation can meet the challenges of the rapidly evolving global capital markets.

About Bob Giffords

Bob Giffords is an independent banking technology analyst based in the UK, who carries out research on the global evolution of the global capital markets. Following a long career in systems development and consulting to the finance sector, he was CTO for a leading bond trading platform in Europe. He is now a frequent speaker at securities industry conferences and writes for various industry journals including The Trade, Waters, Dealing with Technology, Swift Dialogue and Financial World. He can be reached at bob.giffords@btinternet.com.

About StreamBase

StreamBase Systems, Inc, a leader in high-performance Complex Event Processing (CEP), provides software for rapidly building systems that analyze and act on real-time streaming data for instantaneous decision-making. StreamBase's Event Processing Platform(tm) combines a rapid application development environment, an ultra low-latency high-throughput event server, and the broadest connectivity to real-time and historical data and leading EMS/ OMS software platforms. Six of the top ten Wall Street investment banks and three of the top five hedge funds use StreamBase to power mission-critical applications to increase revenue, lower costs, and reduce risk. It is also used by government agencies for highly specialized intelligence work. The company is headquartered in Lexington, Massachusetts with European offices in London. For more information, visit www.streambase.com.

About Vertica

Vertica Systems is the market innovator for high-performance analytic database management systems that run on industry-standard hardware. Co-founded by database pioneer Dr. Michael Stonebraker, Vertica has developed grid-based, column-oriented analytic database technology that lets companies of any size store and query very large databases orders of magnitude faster and more affordably than other solutions. The Vertica Analytic Database is available as software only, as an appliance or online as a cloud computing solution. The technology's unmatched speed, scalability, flexibility and ease of use helps customers like JP Morgan Chase, Verizon, Mozilla, Comcast, Level 3 Communications and Vonage capitalize on business opportunities in real time. Vertica is headquartered in Billerica, Mass. For more information, visit the company's Web site at <http://www.vertica.com>.

Corporate Headquarters
181 Spring Street
Lexington, MA 02421
+1 (866) 787-6227

New York Office
845 Third Avenue, 6th Floor
New York, NY 10022
+1 (866) 787-6227

Virginia Office
11921 Freedom Drive, Suite 550
Reston, VA 20190
+1 (703) 608-6958

European Headquarters
60 Cannon Street
London EC4N 6NP
+44 (0) 20 7002 1095

www.streambase.com